**Career Track Analysis with SQL and Tableau Project**

Exploring Student Enrollments and Completions in Data-Related Career Tracks

**Case Description**

One of the functionalities the 365 company introduced in a 2021 platform release included the option for student enrollment in a career track. The tracks represent an ordinal sequence of courses that eventually lead to obtaining the skills for one of three job titles: data scientist, data analyst, or business analyst.

Completing a career track on the platform is a challenging task. To acquire a corresponding career certificate, a student must pass nine course exams (seven compulsory and two elective courses) and sit for a career track exam encompassing topics from all seven required courses.

In this Career Track Analysis with SQL and Tableau project, you’re tasked with analyzing the career track enrollments and achievements of 365’s students. You’ll first need to retrieve the necessary information from an SQL database. Afterward, you’ll feed this information to Tableau, visualize the results, and finally interpret them.

**Extracting the Data with SQL**

In this Career Track Analysis with SQL and Tableau project, you’re tasked with analyzing the career track enrollments and completions of 365’s students. You’ll first need to retrieve the necessary information from an SQL database. Afterward, you’ll feed this information to Tableau and visualize the results.

Study the sql\_and\_tableau database, consisting of the following tables:

* career\_track\_info
  + track\_id – the unique identification of a track, which serves as the primary key to the table
  + track\_name – the name of the track
* career\_track\_student\_enrollments
  + student\_id – the unique identification of a student
  + track\_id – the unique identification of a track. Together with the previous column, they make up the primary key to the table—i.e., each student can enroll in a specific track only once. But a student can enroll in more than one career track.
  + date\_enrolled – the date the student enrolled in the track. A student can enroll in more than one career track.
  + date\_completed – the date the student completed the track. If the track is not completed, the field is NULL.

Load the database from the sql\_and\_tableau.sql file. Using the career\_track\_info and career\_track\_student\_enrollments tables in the database, select a dataset containing the following columns:

* student\_track\_id – this serves as an identification for each student-track pair. That is, each row of the resulting table should be uniquely labeled. Do this in an ordinal manner.  
  **Hint:***Research how the ROW\_NUMBER function with the OVER clause work.*
* student\_id – the unique identification of a student
* track\_name – the name of the track
* date\_enrolled – the date the student enrolled in the track
* track\_completed
  + 0 – the track is not completed (the completion date is NULL)
  + 1 – the track is completed (the completion date is not NULL)
* days\_for\_completion – the difference in days between the completion date and the enrollment date  
  **Hint:***Research how the DATEDIFF function works.*
* completion\_bucket – the bucket a student falls into based on the number of days it took them to complete a track (if they have done so). The buckets are as follows:
  + Same day – the days\_for\_completion field is equal to 0
  + 1 to 7 days – the days\_for\_completion field is between 1 and 7 days, inclusive
  + 8 to 30 days – the days\_for\_completion field is between 8 and 30 days, inclusive
  + 31 to 60 days – the days\_for\_completion field is between 31 and 60 days, inclusive
  + 61 to 90 days – the days\_for\_completion field is between 61 and 90 days, inclusive
  + 91 to 365 days – the days\_for\_completion field is between 91 and 365 days, inclusive
  + 366+ days – the days\_for\_completion field is more extensive than 365 days

Export the obtained result as a CSV file called career\_track\_completions.csv.

To complete the task, follow these steps.

1. **Subquery:** Begin with a subquery that selects columns from joined tables. You want to join two tables (career\_track\_student\_enrollments and career\_track\_info) based on a common column, track\_id.

SELECT

...

FROM

career\_track\_student\_enrollments e

JOIN

career\_track\_info i ON ...;

1. In the field list of your subquery, you want to select several columns: student\_id, track\_name, date\_enrolled, and date\_completed. You should also create three columns that can’t be found in the database: student\_track\_id, track\_completed, and days\_for\_completion.

SELECT

... AS student\_track\_id,

# Selecting the four columns,

... AS track\_completed,

... AS days\_for\_completion

FROM

career\_track\_student\_enrollments e

JOIN

career\_track\_info i ON ...;

1. To create the student\_track\_id column, use the ROW\_NUMBER function with the OVER clause. This assigns unique row numbers to each row in the result set. The row numbers are ordered by student\_id and track\_name in descending order. The function should look like the following:

ROW\_NUMBER() OVER (...) AS student\_track\_id,

1. To create the track\_complete column, check whether date\_completed is NULL. In SQL, you can use the IF function to do this.

IF(...) AS track\_completed

1. To calculate the days\_for\_completion column, you can use the DATEDIFF function in SQL, which calculates the difference between two dates.

DATEDIFF(...) AS days\_for\_completion

1. This entire subquery will act as a new dataset in your main query. So, enclose it in parentheses and give it an alias, say a.

(

...

JOIN

career\_track\_info i ON ...) a;

1. **Main Query:** Select the columns you want using a SELECT statement in the main query. You will also create a new column, completion\_bucket.

SELECT

# Select all columns from the subquery,

... AS completion\_bucket

FROM

( # The subquery ) a;

1. To create the completion\_bucket column, use the SQL CASE statement. The CASE statement allows you to perform conditional logic in SQL; it behaves similarly to if-else conditions in programming. Each WHEN clause within the CASE statement will check if days\_for\_completion falls within a specific range, and the THEN clause will assign the appropriate bucket label. After writing all the WHEN-THEN pairs, finish the CASE statement with the END keyword, and the alias completion\_bucket.

SELECT

# Select all columns from the subquery,

CASE

WHEN days\_for\_completion = 0 THEN 'Same day'

...

END AS completion\_bucket

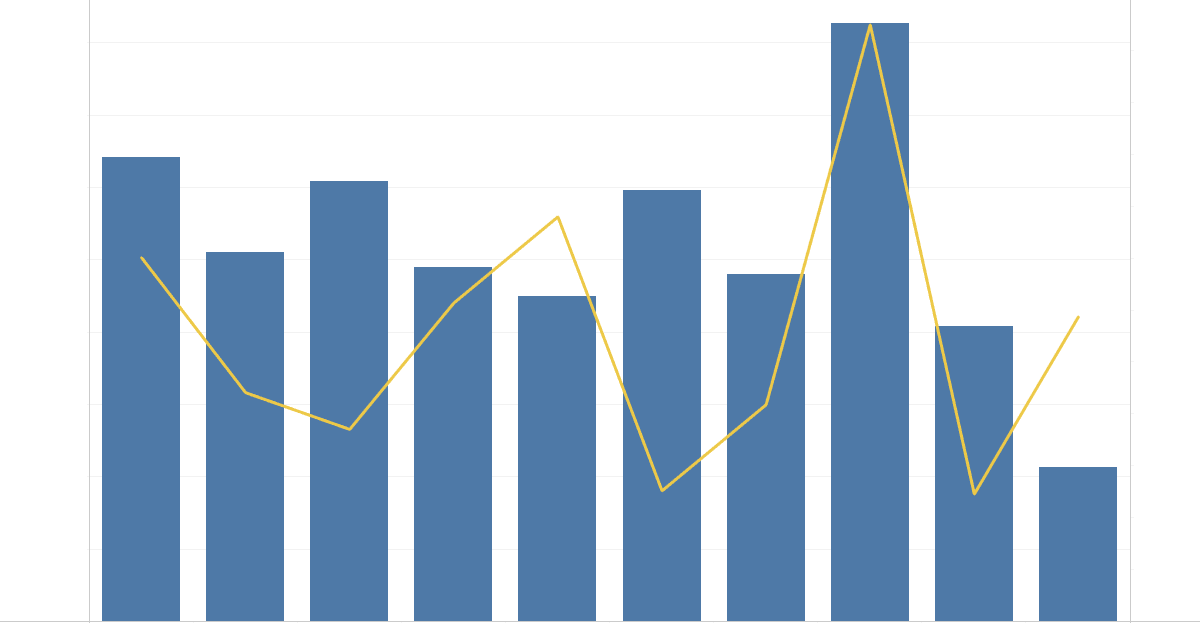
FROM

( # The subquery ) a;

**Creating a Combo Chart in Tableau**

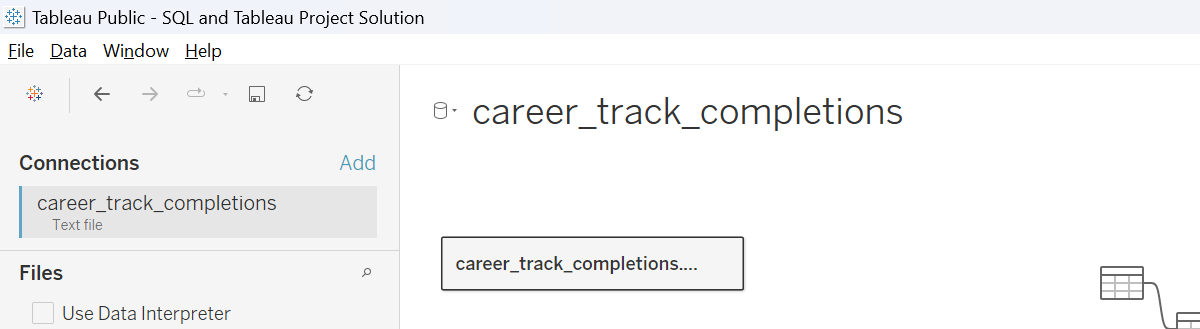
For this part of the project, you’d need to open Tableau and feed it the data source you retrieved in the previous part. Your task is to create a combo chart—a bar chart overlaid with a line chart. Each bar should represent a given month, while their height should represent the number of track enrollments that month (This is **not** the same as the number of students who’ve enrolled in a track*.*)

The line chart should, in turn, display the fraction of track completions that month as a percentage of the number of enrollments. Add to the chart a filter that sifts out the different career tracks.

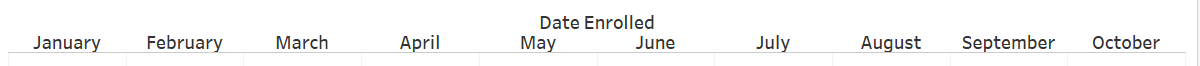


Use the following directions to carry out the task.

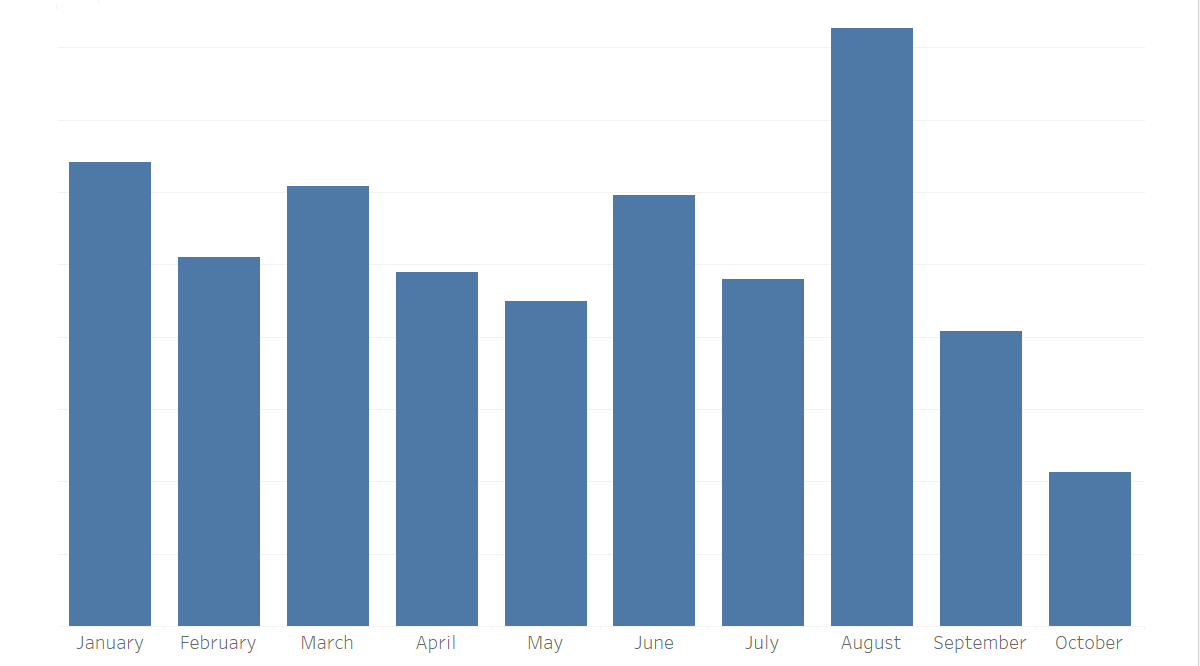
1. **Connect to the data source:** After opening Tableau, the first step is to connect to your data source.



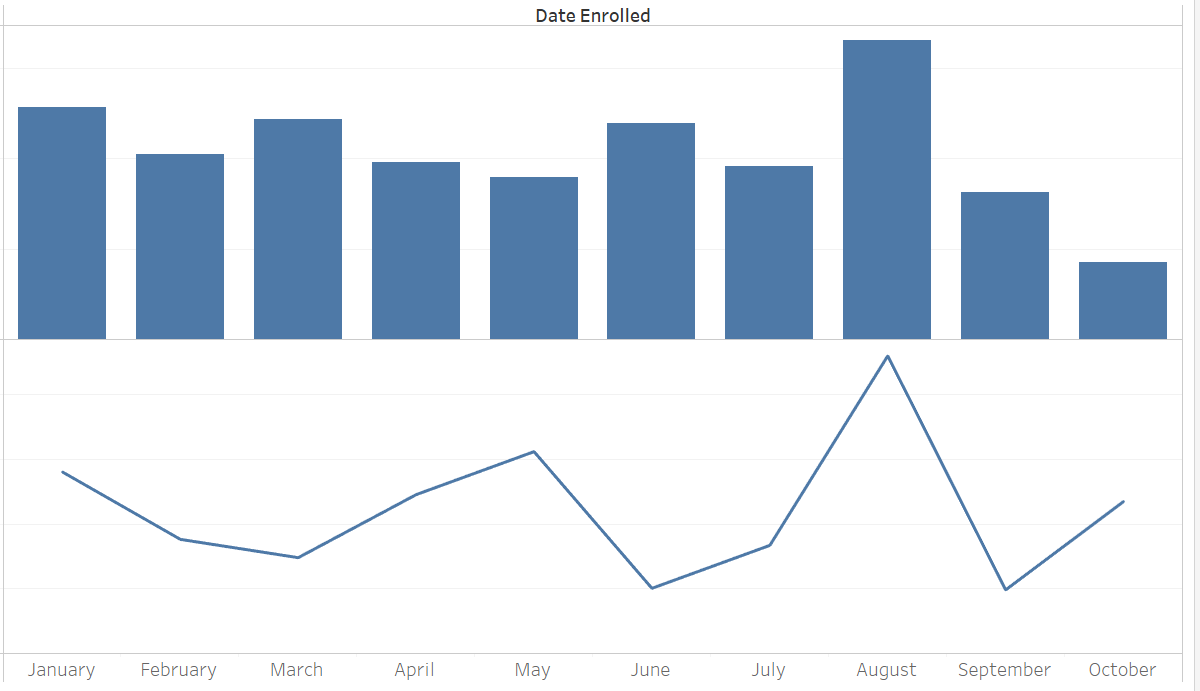
1. **Create the bar chart:**
   * Drag the Date Enrolled dimension to the Columns shelf and change the function from YEAR to MONTH. This will create an axis for the months.



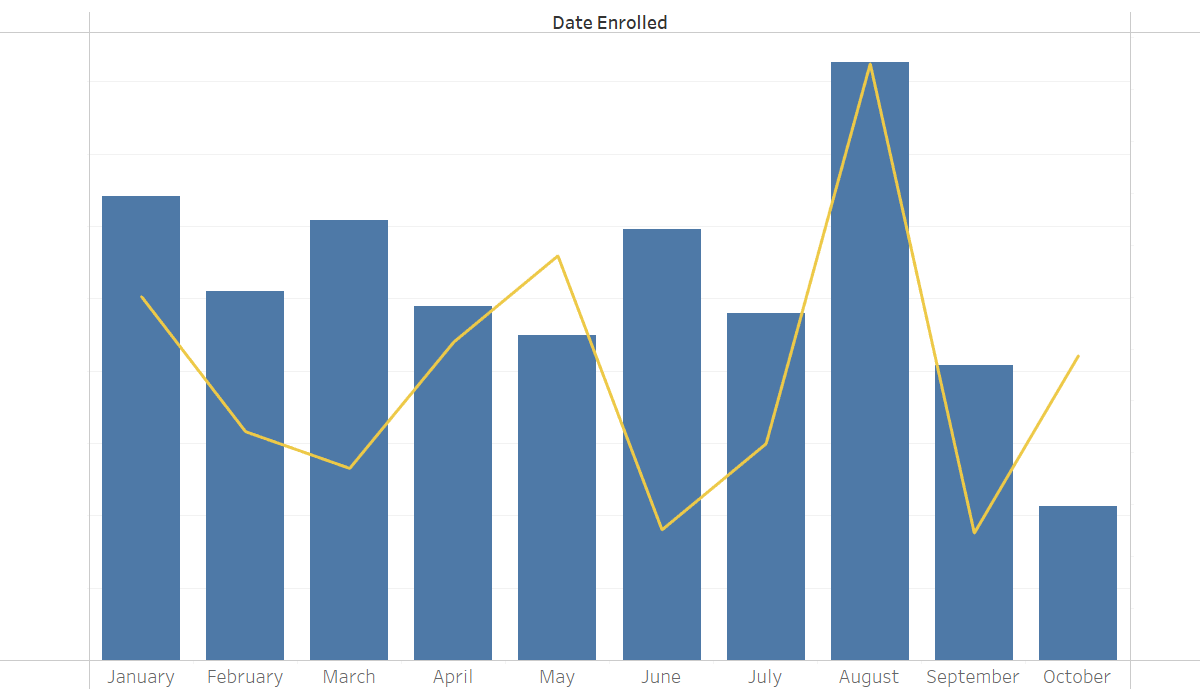
* + Ensure the Student Track ID field is under Measures. Drag it to the Rows shelf and change the function to COUNT. This will create the bars with their heights proportional to the number of enrollments in each month. If the automatic chart is not a bar chart, change the type accordingly. Additionally, ensure all months from January to October are displayed on the graph.



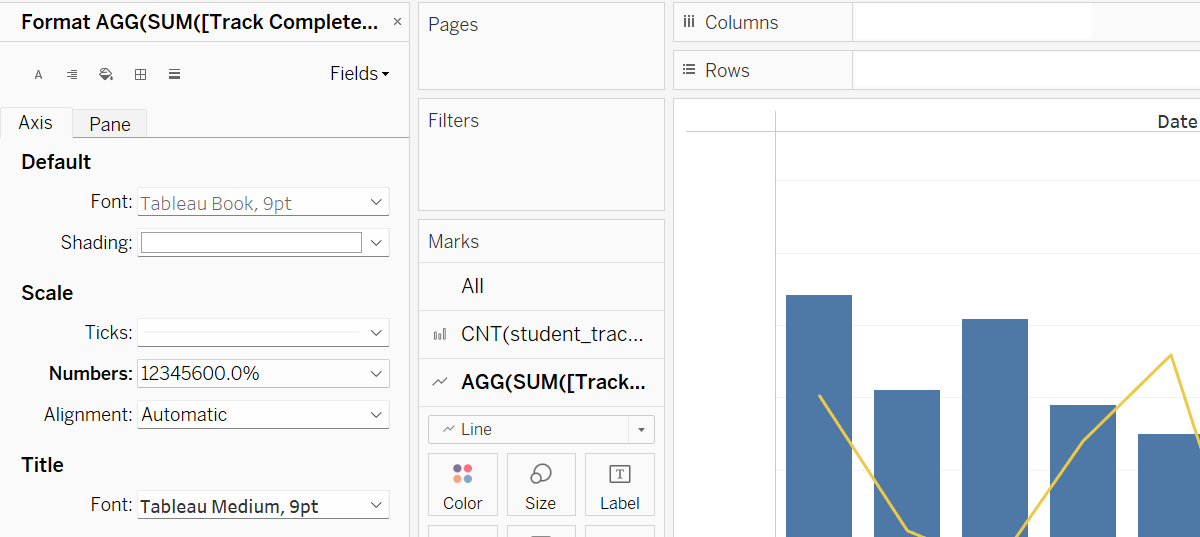
1. **Create the line chart:** Now, drag the Track Completed measure to the right of the first measure on the Rows shelf, adding another axis and a second chart. Ensure that second chart is of a line type. Leave the function as SUM. Divide the resulting number by the number of track enrollments, i.e., COUNT([student\_track\_id]).



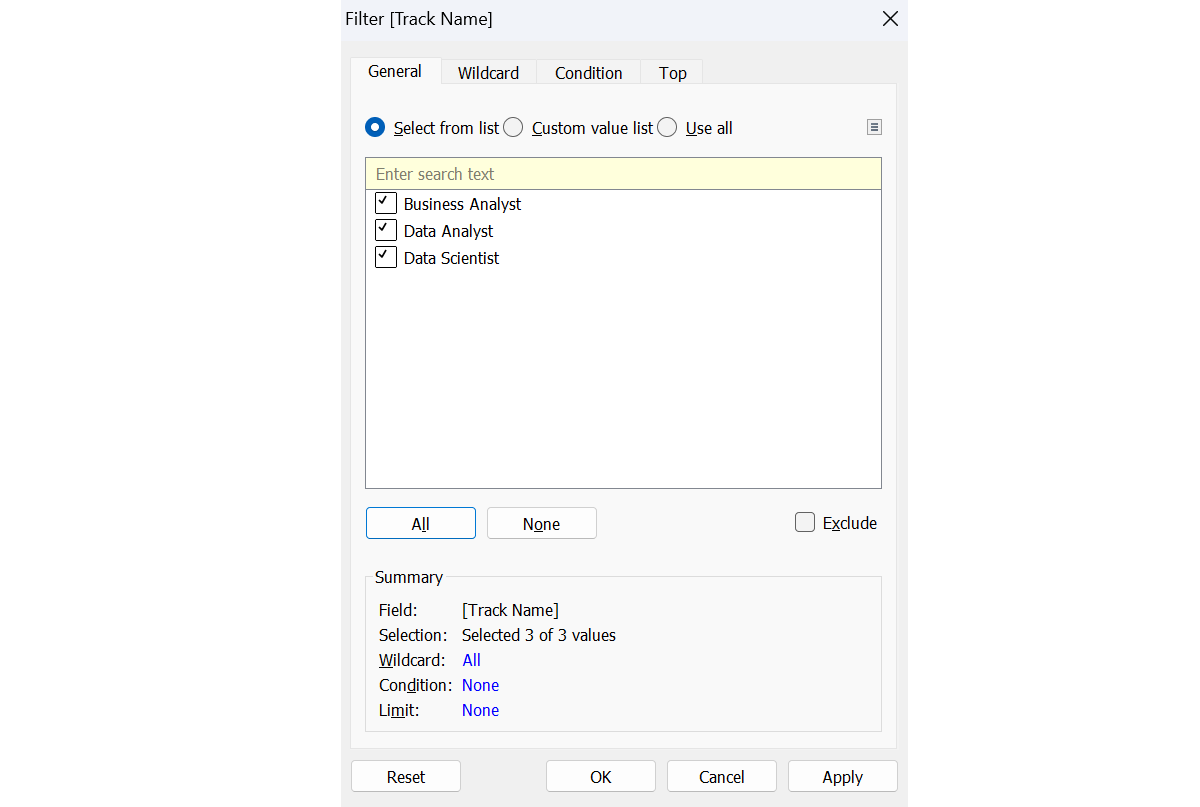
1. **Dual axis:** Now, you have two charts—one above the other. To overlay them, right-click on the bottom yy-axis and select Dual Axis. Alternatively, right-click on the second item in the Rows shelf and select Dual Axis.



1. **Change the y-axis:** Make sure the yy-axis corresponding to the line chart represents the percentage values. This can be done by right-clicking on the axis, choosing Format, navigating to the Axis pane, then Scale, Numbers, and choosing Percentage. Decrease the decimal places to 1.

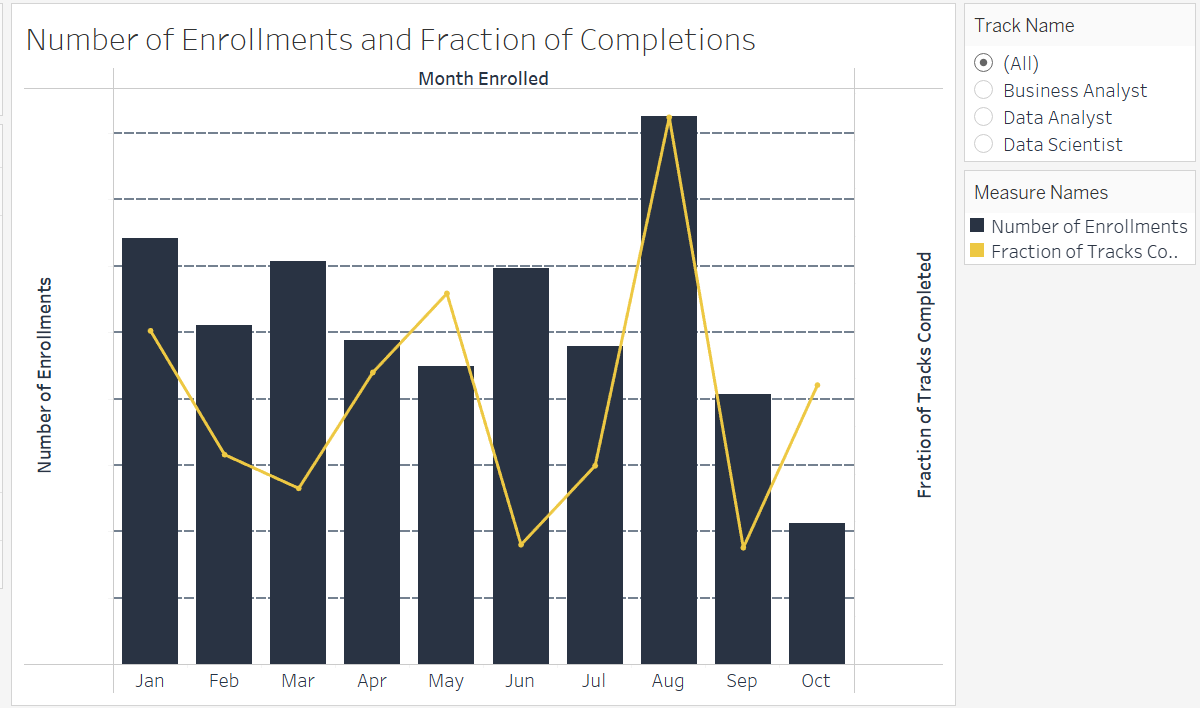


1. **Create the filter:** Drag the Track Name Dimension to the Filters shelf and select All. Show the filter and make sure it works as expected.



1. **Style, title, and labels (Optional):** Research how to adjust the color, style, and size of the elements in your chart. You can add a legend to your view to denote which mark type corresponds to which measure. Finally, you can add title and tooltips to your chart and labels to your axes.

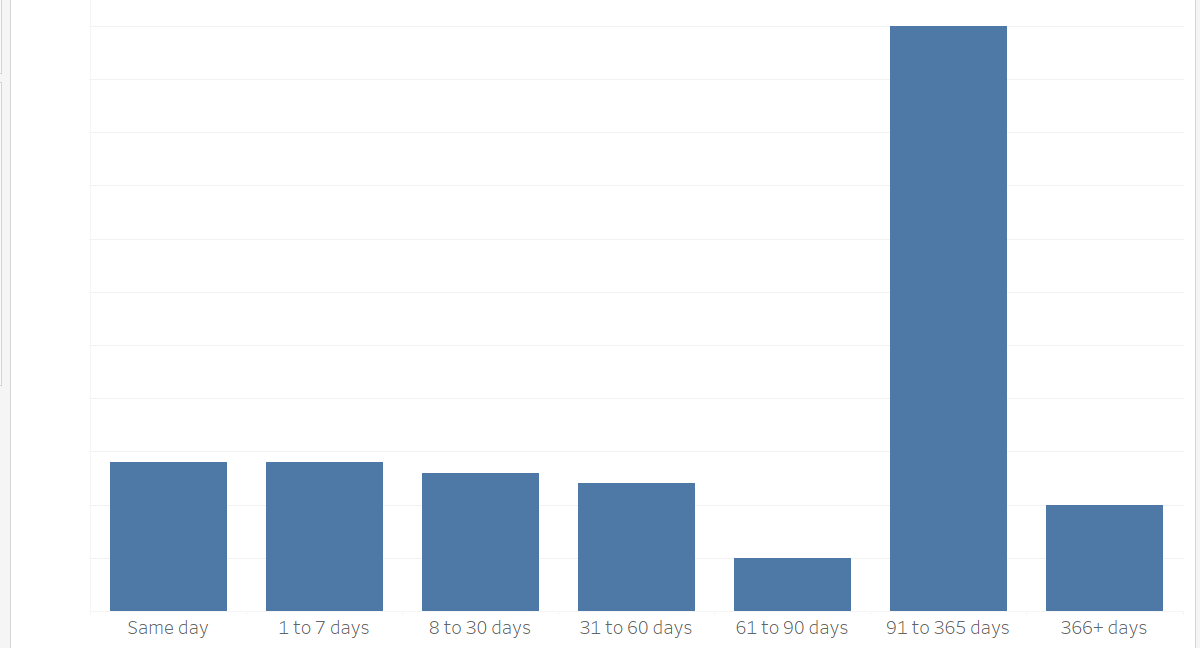
Below is an example you can draw inspiration from.



**Creating a Bar Chart in Tableau**

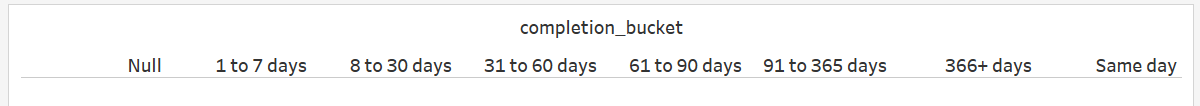
Congratulations on bulding the combo chart! It’s now time to create a new worksheet containing a bar chart with each bar representing a different completion bucket and their height corresponding to the number of track enrollments completed in the respective time. Add a filter that sifts out the various career tracks. Make sure to remove the NULL bar. Order the bars such that—from left to right—they are ordered in the following way:

* Same day
* 1 to 7 days
* 8 to 30 days
* 31 to 60 days
* 61 to 90 days
* 91 to 365 days
* 366+ days

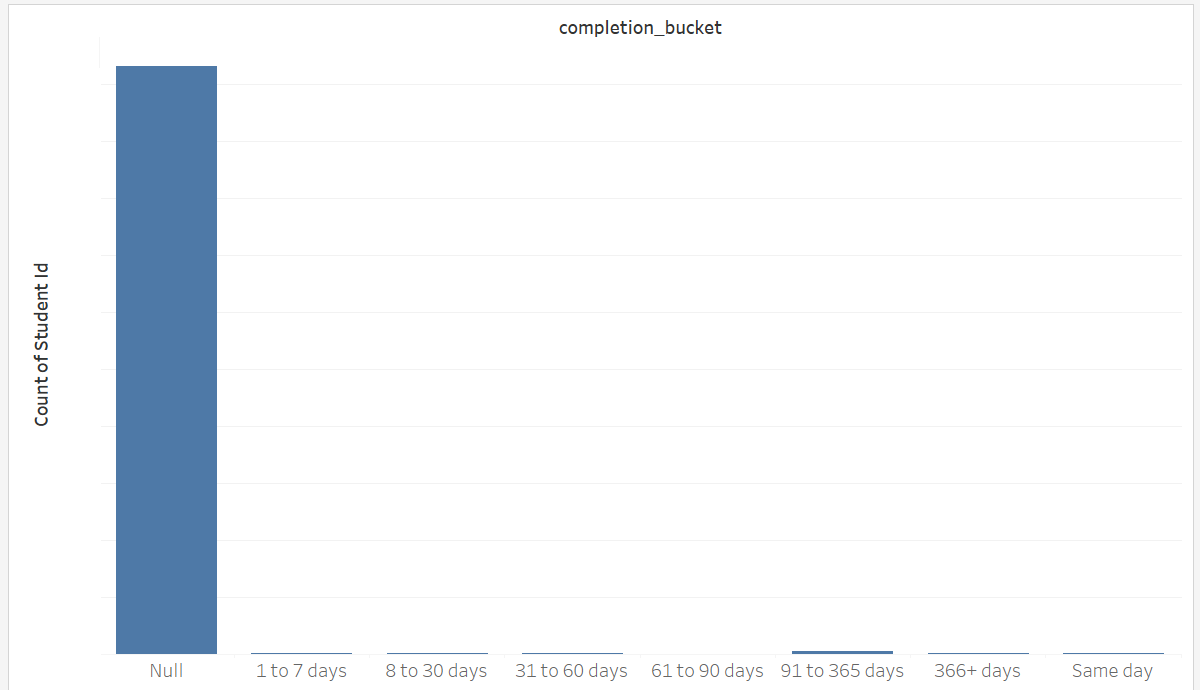


Follow the steps below to perform the task.

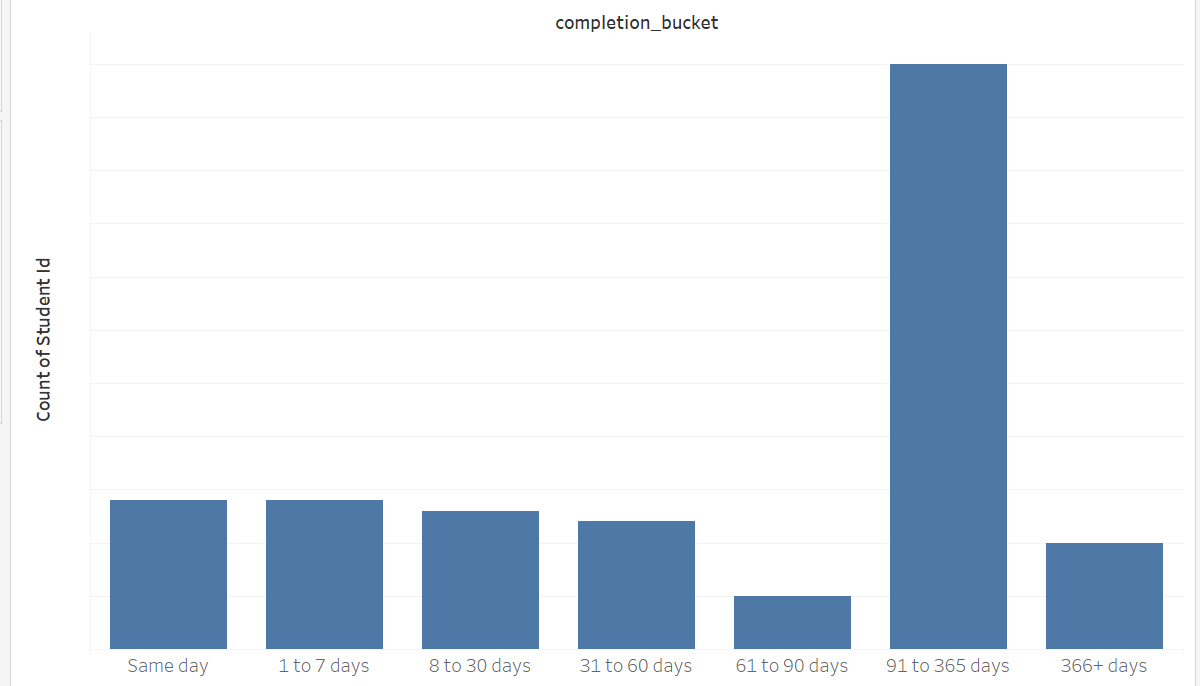
1. **Create the bar chart:**
   * Drag the Completion Bucket Dimension to the Columns shelf.



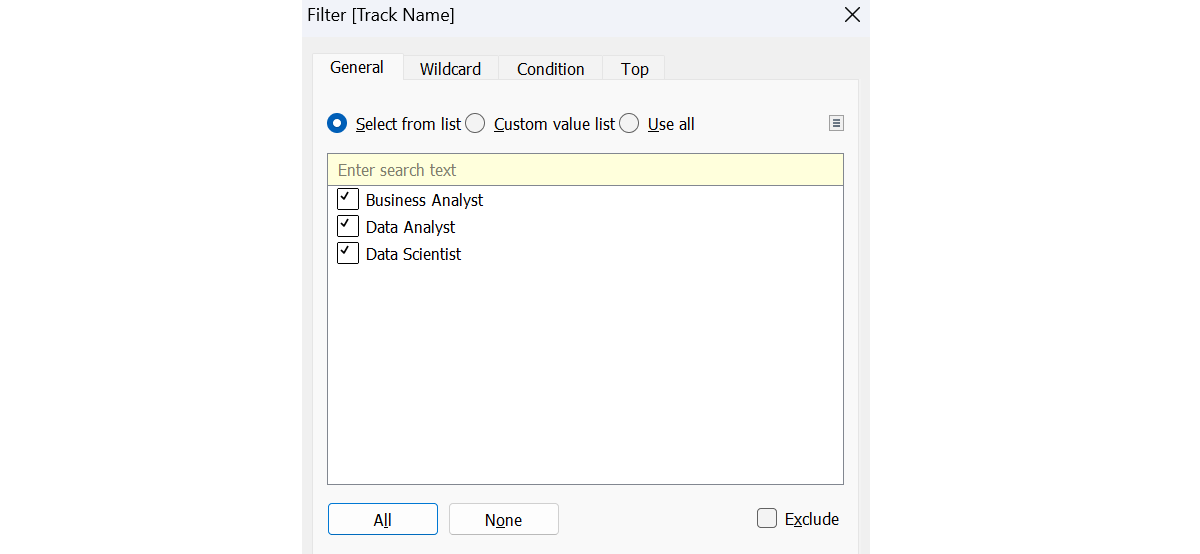
* + Ensure the Student Track ID field is under Measures. Drag it to the Rows shelf and find its count. Make sure the type of chart is a bar.



* + Remove the NULL-values bar. Rearrange the bars if necessary.

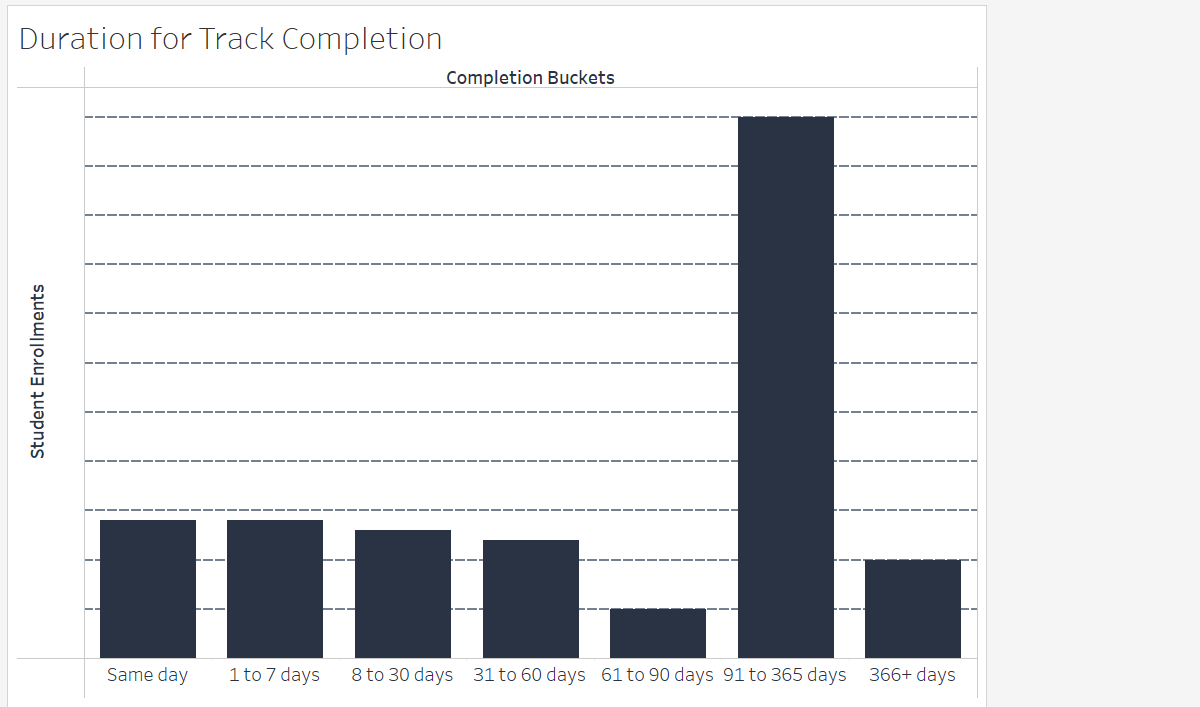


1. **Create the filter:** Drag the Track Name Dimension to the Filters shelf and select All. Show the filter and make sure it works as expected.



1. **Style, title, and labels (Optional):** Research how to adjust the color, style, and size of the elements in your chart. You can add a legend to your view to denote which mark type corresponds to which measure. Finally, you can add title and tooltips to your chart and labels to your axes.

Below is an example you can draw inspiration from.



**Interpreting the Results**

Congratulations on completing the practical part of the project! Now, analyzing the study results and drawing meaningful conclusions is essential.

Consider the Tableau visualizations you created and answer the following questions:

1. What is the number of enrolled students monthly? Which is the month with the most enrollments? Speculate about the reason for the increased numbers.
2. Which career track do students enroll most in?
3. What is the career track completion rate? Can you say if it’s increasing, decreasing, or staying constant with time?
4. How long does it typically take students to complete a career track? What type of subscription is most suitable for students who aim to complete a career track: monthly, quarterly, or annual?
5. What advice and suggestions for improvement would you give the 365 team to boost engagement, increase the track completion rate, and motivate students to learn more consistently?

**Quiz**

**Question 1:**

**Study your solution to Extracting the Data with SQL. How many days did it take the student with the most extended completion period to complete a career track?**

**482 days**

**Question 2:**

**Study your solution to Extracting the Data with SQL. Referring to the student in the previous question, what career track did the student complete**

**Data Analyst**

**Question 3:**

**Study your solution to Extracting the Data with SQL. How many track completions are there in total?**

**123**

**Question 4:**

**Study your solution to Creating a Combo Chart in Tableau. Consider all career tracks and the month with the highest fraction of track completions. Roughly how many people have enrolled in a career track in that month?**

**1,650**

**Question 5:**

**Study your solution to Creating a Combo Chart in Tableau. Which career track has the highest fraction of track completions in May?**

**Business Analyst**

**Question 6:**

**Study your solution to Creating a Combo Chart in Tableau. Which career track has the lowest fraction of track completions in January?**

**Data Scientist**

**Question 7:**

**Study your solution to Creating a Bar Chart in Tableau. When considering the enrollments in the Data Scientist career track, which of the following buckets contains the smallest number of representatives?**

**days31 to 60**

**Question 8:**

**Study your solution to Creating a Bar Chart in Tableau. When considering the enrollments in the Data Analyst career track, which of the following buckets contains the highest number of people?**

**91 to 365 days**